Geochemistry Geophysics Geosystems G3
AN ELECTRONIC JOURNAL OF THE EARTH SCIENCES
Published by AGU and the Geochemical Society

Technical Brief
Volume 7, Number 9
28 September 2006
Q09013, doi:10.1029/2006GC001272
ISSN: 1525-2027

Wavelet analysis of the seismograms of the 2004 Sumatra-Andaman earthquake and its application to tsunami early warning

Oliver G. Lockwood
Pembroke College, Cambridge University, CB2 1RF Cambridge, UK (oliver.lockwood@cantab.net)

Hiroo Kanamori
Seismological Laboratory, California Institute of Technology, 1200 East California Boulevard, Pasadena, California 91125, USA (hiroo@seismo.gps.caltech.edu)

[1] We applied the wavelet transform in an attempt to detect long-period components early in a seismogram. We analyzed the displacement seismograms of the 26 December 2004 Sumatra-Andaman earthquake (M_w = 9.2) and the 28 March 2005 Nias earthquake (M_w = 8.7). Wavelet analysis is able to clearly distinguish the amplitudes of the long-period W phase between the seismograms of the two earthquakes before the S wave reaches the station. It shows that the 2004 earthquake generates a W phase of significantly greater amplitude. This facility has potential application to the rapid identification of truly great earthquakes with high tsunami potential.

Components: 3392 words, 8 figures.

Keywords: tsunami warning; wavelet; W phase; scalogram.

Index Terms: 7215 Seismology: Earthquake source observations (1240); 7255 Seismology: Surface waves and free oscillations.

Received 7 February 2006; **Revised** 17 June 2006; **Accepted** 18 July 2006; **Published** 28 September 2006.

Lockwood, O. G., and H. Kanamori (2006), Wavelet analysis of the seismograms of the 2004 Sumatra-Andaman earthquake and its application to tsunami early warning, *Geochem. Geophys. Geosyst.*, *7*, Q09013, doi:10.1029/2006GC001272.

1. Introduction

[2] The M_w = 9.2 Sumatra-Andaman earthquake on 26 December 2004 caused unprecedented damage to many countries around the Indian Ocean. Had an effective early warning system existed, many casualties could have been averted; if not in nearby areas such as Sumatra itself, then certainly in more distant locations such as Sri Lanka, Thailand where a matter of hours passed between generation of the tsunami and its local arrival. To establish an effective tsunami warning system, a comprehensive program is necessary. This should include monitoring seismic waves, crustal deformations, water waves, infrastructure for information transfer and logistics, and education and training of residents. Direct water-wave monitoring of tsunami, such as exists in parts of the Pacific [*National Oceanic and Atmospheric Administration*, 2006], is useful for detecting tsunami in the open sea, but it must be located in the area where warnings are needed. At present, no such system exists in the Indian Ocean. Also, to build and maintain

such a system requires substantial amounts of resources.

[3] Seismic tsunami warning systems have a long history in the U.S. and Japan and several systems are now in operation mainly for the Pacific (see *Bryant* [2001] for a recent review). With the recent availability of high-quality global seismic data, seismologists can rapidly determine many of the important physical characteristics of earthquakes, yet it took seismologists hours to recognize how large the 26 December event really was, partly because the present global observation systems are not specifically designed for such extremely large events [*Kerr*, 2005]. The 2004 Sumatra-Andaman earthquake motivated seismologists to investigate new methods [*Menke and Levin*, 2005; *Lomax and Michelini*, 2005; *Bormann and Wylegalia*, 2005] for identifying the tsunami potential of an earthquake.

[4] Here, to complement these studies we suggest a simple seismological method that can rapidly distinguish truly great earthquakes from large earthquakes using very long-period waves which arrive before the S phase. The use of long-period seismic waves for tsunami warning purposes have been discussed in the literature [e.g., *Kanamori and Given*, 1983; *Talandier and Okal*, 1989; *Okal et al.*, 1991], and here we focus on exceptionally large earthquakes (i.e., $M_w \geq 9$) like the 2004 Sumatra-Andaman earthquake for which the present global system did not work well.

[5] For illustration purposes, we investigate the two recent Sumatran earthquakes; the aforementioned December 2004 Sumatra-Andaman earthquake, and the $M_w = 8.7$ Nias earthquake on 28 March 2005. The main reason for this was the close proximity of the two epicenters, which results in similar path effects between the source of either event and any given seismic monitoring station in the world. This simplifies comparison of data for the two events, such that variation between the records at a given station is almost entirely due to differences in the events themselves. Although also large in M_w, the 2005 Nias earthquake did not generate as significant a tsunami. While the tsunamigenic potential of an earthquake depends not only on its size but also its location and depth, most devastating widespread tsunamis are generated by those events with $M_w \geq 9$.

[6] Our analysis concerned itself with detecting and quantifying characteristic features in the seismic data which have the potential to indicate the presence or lack of a tsunami. We studied data from a selection of seismic stations at varying azimuth from the epicenter, as shown in Figure 1.

2. Real-Time Processing and Recursive Deconvolution

[7] We analyze ground-motion displacement data. Typically the seismometer output is deconvolved with the instrument's frequency response over the whole event's duration. However, since it requires knowledge of the future, this method is rendered useless for early warning. For real-time application, we converted the raw seismic data into ground-velocity by using a recursive deconvolution described by *Zhu* [2003], and then integrating it with a high-pass time domain Butterworth filter to reduce the effects of instrument drift. The corner period of the filter was set to be 2000 s, twice the period of interest, in line with standard signal processing practice. This method yielded ground-displacement records similar to those produced in the traditional manner.

3. Seismic Records

[8] Our analysis was performed on vertical ground-displacement data. To illustrate our method, we use the records from the station OBN (Obninsk, Russia). We aligned all records such that the P wave arrival occurs 1000 s into the signal duration. Figure 2a shows the displacement seismogram of the 2004 Sumatra-Andaman earthquake on which we mark the arrival times of P, PP, and S phases. The long-period (i.e., low-frequency) ramping W phase [*Kanamori*, 1993] is evident between the arrivals of the PP and S waves. W phase can be interpreted as superposition of overtones of long-period Rayleigh waves or superposition of multiply-reflected phases like PP, PPP, etc. It travels with a group velocity faster than the S wave, and can be effectively used for rapid tsunami warning purposes. Figure 2c shows the displacement record of the 2005 Nias earthquake with the same scale as Figure 2a. Although the W phase can be identified, it is much smaller than that in Figure 2a. The tsunami-generating capability of an earthquake depends on the volume of the displaced water due to seafloor deformation, provided that this deformation occurs sufficiently quickly so that the water cannot flow away from the source. The volume of the displaced water is approximately proportional to the seismic moment, which determines the source spectral amplitude at periods longer than the corner period [*Aki*, 1972].

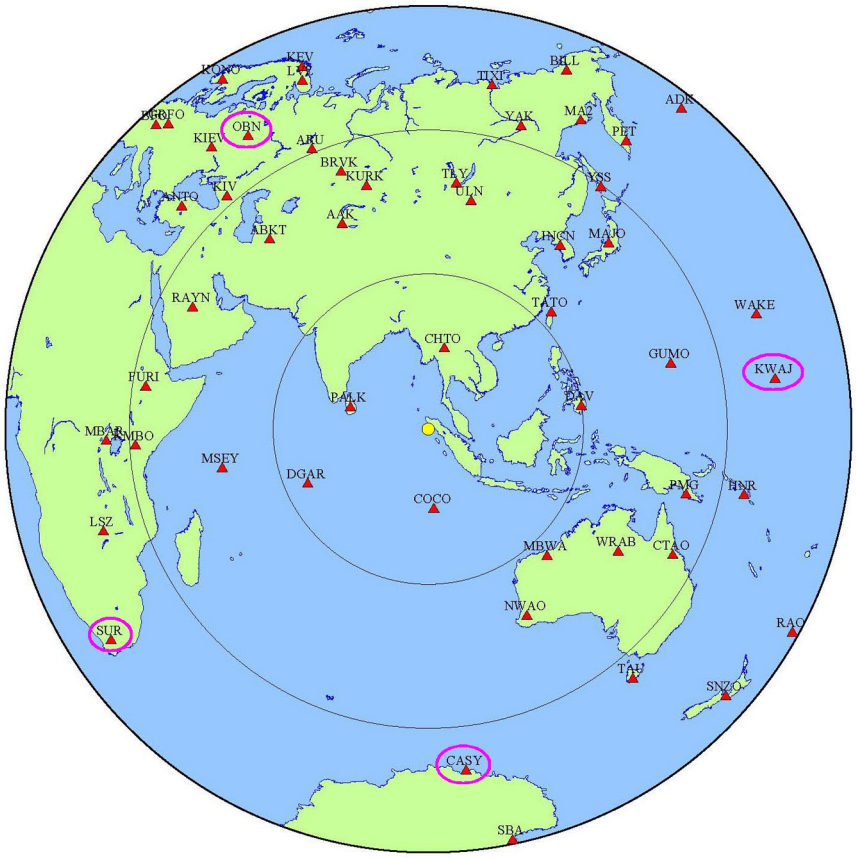

Figure 1. Seismic stations of the IRIS Global Seismic Network. The stations used in our analysis are circled.

Since the corner period of great earthquakes is longer than 100 s, the amplitude of seismic waves with a period longer than 100 s is most relevant to tsunami generation. Thus the fundamental challenge faced is to find a way to identify the long-period W phase and measure its amplitude quickly.

4. Spectral Analysis

[9] The use of the Fourier transform (FT) enables a signal to be decomposed into individual frequency components, i.e., a set of infinite sinusoids of various period and amplitudes, which when summed back together exactly reproduce the original waveform. While this accurately shows what frequencies are present in a signal, it does not tell us when in the signal's duration each frequency occurs [*Cohen*, 1989].

[10] The windowed STFT (Short-Time Fourier Transform) has been often used to produce a picture of a signal's changing frequency components through its duration. This works by looking at sections of time or "windows" of the signal, and performing standard Fourier analysis on each time section in turn; as displayed in Figure 3. In this way, a so-called "spectrogram" can be formed, with color intensity showing magnitude of coefficients as a function of frequency and time. This works well in signals whose spectral content changes relatively slowly in comparison to its frequencies. Unfortunately, the STFT suffers from a major drawback [*Huerta-Lopez et al.*, 2000], due to Heisenberg's Uncertainty Principle. If the length of the time window is increased, better frequency resolution can be obtained in the spectrogram, but the process suffers the original problem of losing locality information. Conversely, shortening the time window improves the time resolution but frequency information is lost, especially for low-frequency waves. This is an important issue as our interest lies with the long-period components.

5. Wavelet Transform

[11] Instead of separating into infinite sinusoids, the Wavelet Transform (WT) decomposes a signal into a series of finite-duration waves; hence the name "wavelet." Hence wavelet techniques are

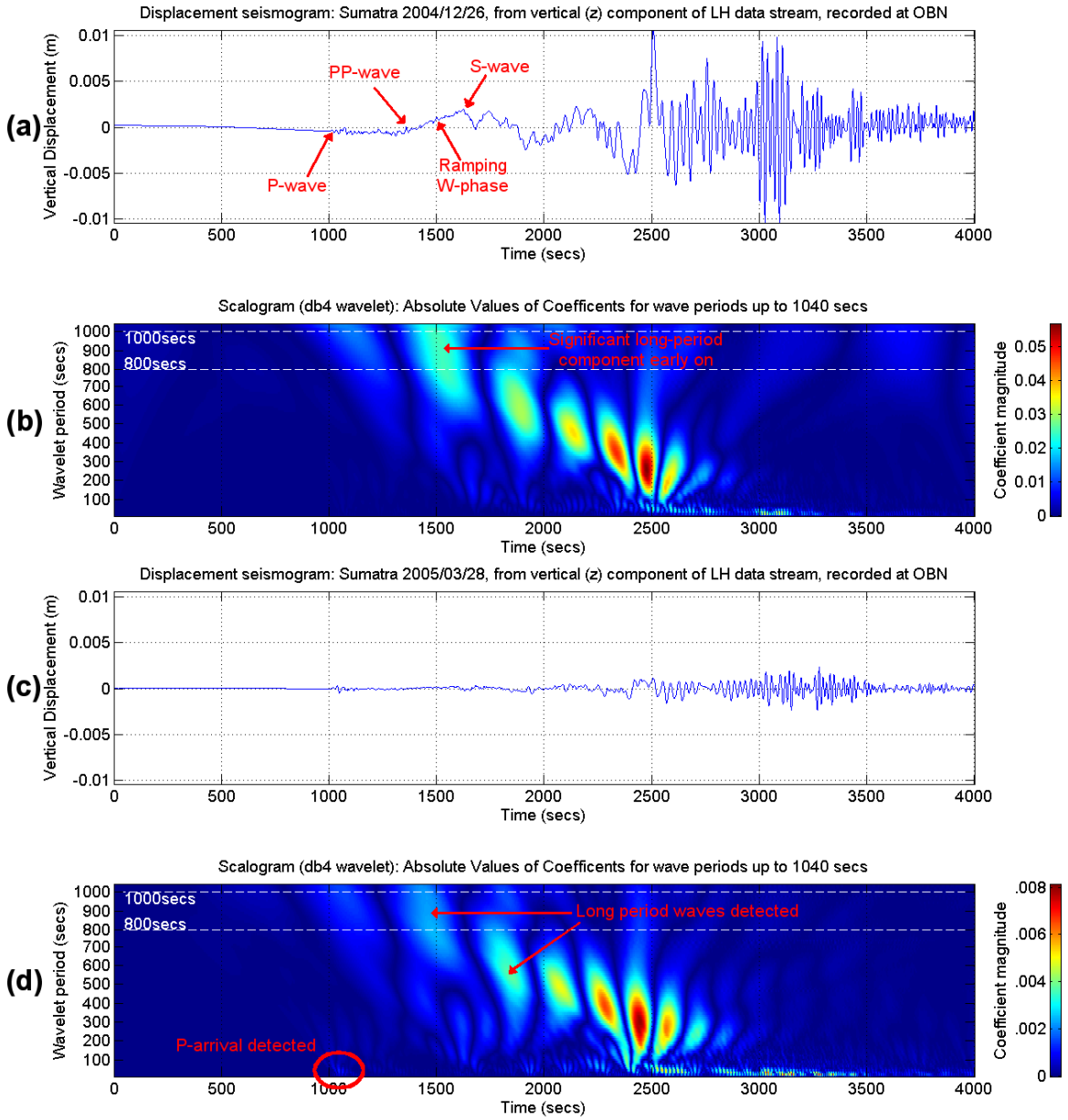

Figure 2. (a) Displacement seismogram of the 2004 Sumatra-Andaman earthquake recorded at OBN. (b) Scalogram of Figure 2a. (c) Displacement seismogram of the 2005 Nias earthquake recorded at OBN. (d) Scalogram of Figure 2c. Notice that the color scale differs greatly from that in Figure 2b. The 2005 event contains much smaller components than the 2004 event, but the picture has been scaled up to make the features visible. A similar shaped, though subtly different, pattern can be observed.

intrinsically better suited to the analysis of transient signals such as seismograms. The wavelet is scaled, i.e., stretched in time, in order to correlate with different frequency components. The associated window size is therefore variable depending on the wavelet scaling that is being looked at [*Chakraborty and Okaya*, 1995]. A higher wavelet scaling corresponds to a longer time period and hence a lower frequency.

[12] The wavelet analysis discussed in this paper employed application of the Continuous Wavelet Transform (CWT) using the Daubechies-4 ("db4") wavelet. This appeared effective at identifying features in seismic data. The low-level calculations were all performed using software within the standard wavelet toolbox for MATLAB 6.5. The CWT is defined as the sum over all time of the signal $f(t)$ multiplied by scaled, shifted versions

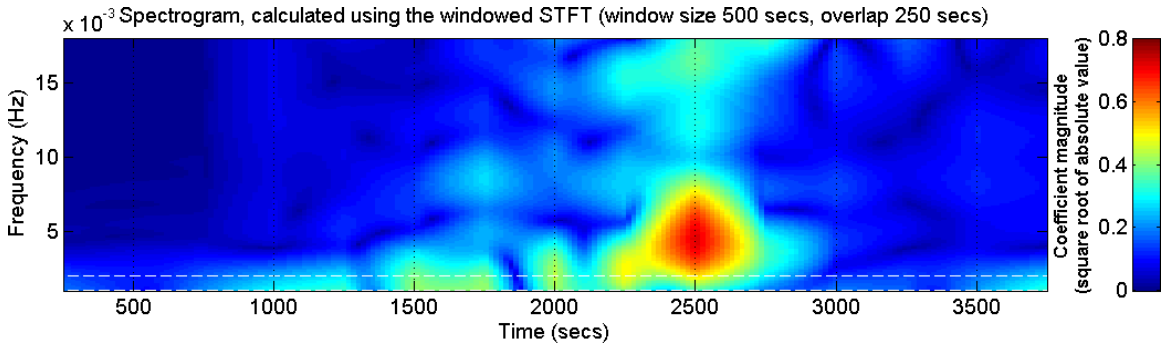

Figure 3. The STFT spectrogram of the displacement seismogram for the 2004 Sumatra-Andaman earthquake.

of the wavelet function ψ. The wavelet transform C is then given by

$$C(scale, position) = \int_{-\infty}^{\infty} f(t)\psi(scale, position, t)dt$$

[13] *Position* refers to position in time with respect to the signal $f(t)$. Since our signals are finite, the integration limits become the size of the signal. For more details on the mathematics involved in the calculation of the wavelet transform, see *Misiti et al.* [1996].

[14] A diagram which displays the wavelet scale as a function of time is called a "scalogram." Figures 2b and 2d show the scalogram for the 2004 and 2005 events, respectively. Color intensity at any point in the picture corresponds to the coefficient magnitude of a wavelet with a particular period at a particular point of the time series. The *y axis* has been translated from wavelet scale into corresponding wavelet time period. (Each particular wavelet has a corresponding "center frequency," which is the frequency for which a sinusoidal approximation gives the best fit. In the case of the Daubechies-4 wavelet, this is 0.7143 Hz.) The long-period component arriving at about 1500 s is much stronger in the 2004 Sumatra-Andaman earthquake than in the 2005 Nias earthquake. This difference is noticeable not only in absolute terms, but also in relative terms compared to the strongest component in the whole signal's duration. (As a result of the very large rupture length of the 2004 event, its spectrum is enhanced at long period.) The WT can simultaneously achieve (1) accurate frequency representation for low frequencies and (2) good time resolution for high frequencies.

[15] For tsunami warning purposes, we chose to concentrate on the very long period waves of 800–1000 s period, as the W phase was found to show up most strongly over this period range. The long-period energy comes from a long rupture time (roughly, rupture length divided by rupture speed). In Figures 2b and 2d, the horizontal white dotted lines are plotted at wave periods of 1000 and 800 s. Having determined this to be a key frequency range for tsunami generating potential, we then tracked the maximum wavelet coefficient within this period range. This gives us a single trace, plotting "long-period wavelet coefficient amplitude" against time. Figure 4 shows the plot for the 2004 event. The ripple visible in Figure 4 and in the scalograms (Figures 2b and 2d) is an effect caused by varying phase of the wavelet. A wavelet

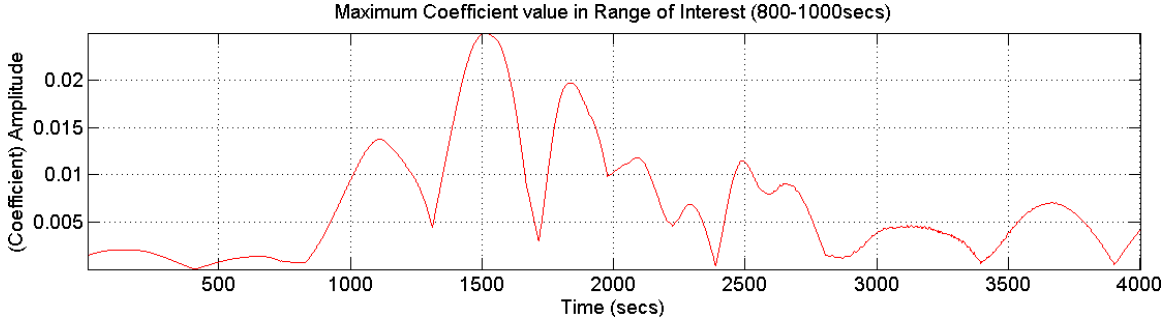

Figure 4. Long-period coefficient amplitude plot for the 2004 Sumatra-Andaman earthquake.

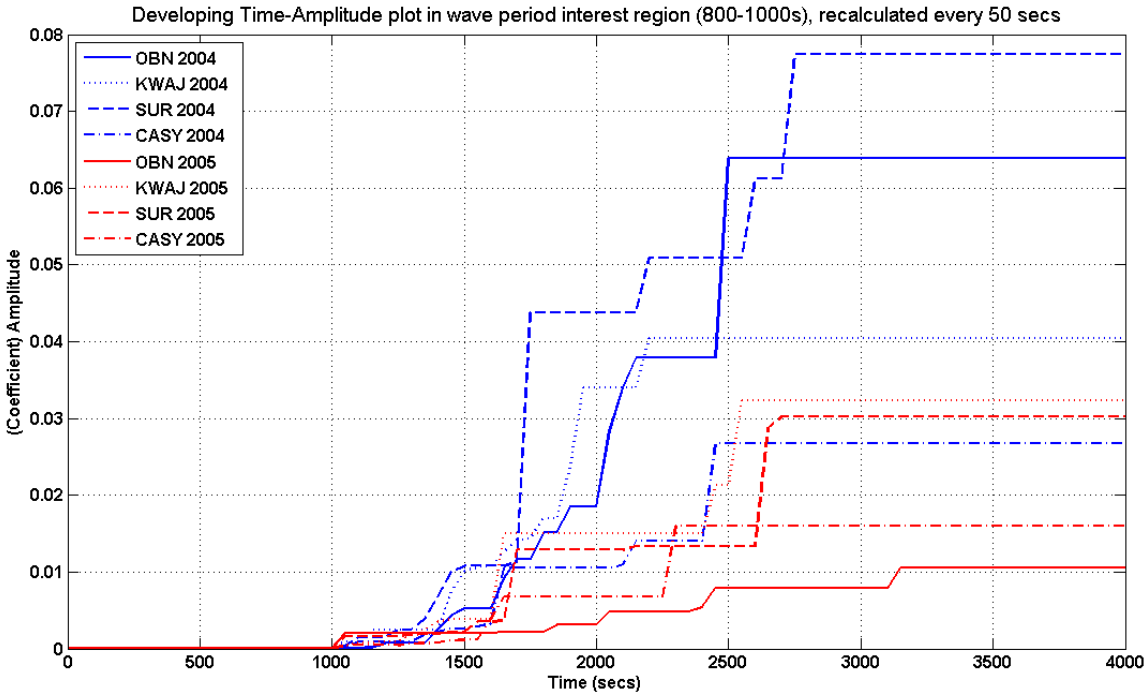

Figure 5. Demonstration of real-time application. Each data point represents the maximum long-period wavelet coefficient measured in a WT which was calculated using only data available up to the marked time point.

is not an infinite sinusoid and therefore, unlike the standard STFT, would not produce a line at a single frequency in the scalogram even if the source were to be a single-frequency pure sinusoid. The WT operates by scaling the wavelet to different periods and then correlating the scaled wavelet along the duration of the signal. As this correlation occurs, the wavelet would go in and out of phase with a pure sinusoid. This oscillation produces the ripple-like effect which is characteristic of the WT, as observed in Figures 2b, 2d, and 4.

6. Real-Time Implementation

[16] In the context of early warning, there are some limitations to the analysis shown so far. While the results indicate that the long-period component is evident within the first 500 s after P wave arrival, the wavelet transform which shows this has been actually calculated across the whole 4000 s of data. The logical next step is to investigate whether these long-period waves can be picked up in real-time as the waveform progresses, or whether they require the full signal to calculate. Unfortunately it is not possible to perform the wavelet transform in a recursive manner as for the deconvolution of the seismogram. A possible solution is to recalculate the WT after a block of new data, say every 50 s. We took the recursively deconvolved ground-displacement data for both events and then repeatedly performed the following steps, with our starting point as before 1000 s before the P wave arrival: (1) adding 50 s of data, (2) recalculating the WT, and (3) extracting the maximum long-period coefficient in the signal. Figure 5 shows the output from this plotted against time. In this graph every data point is based only on information available at the time for which it is plotted. Despite some noise in the signal, we can see that the 2004 event rapidly distinguishes itself as having a much stronger long-period component. This is clearer in the case of some stations than others. The noise is to some extent an artifact caused by the 2-pass (forward and back) Butterworth filtering accompanying the recursive filter on an incomplete data set.

[17] In the above we mainly illustrated our method using the records from the station OBN, but the W phase has a relatively uniform radiation pattern and can be identified equally well at other stations. Figures 6, 7, and 8 show the displacement records and scalograms at other stations, KWAJ (Kwajalein Atoll, Marshall Islands), SUR (Sutherland, South Africa), and CASY (Casey, Antarctica) shown in Figure 1. These stations are all approximately 70° distance from the epicenter, and were chosen for

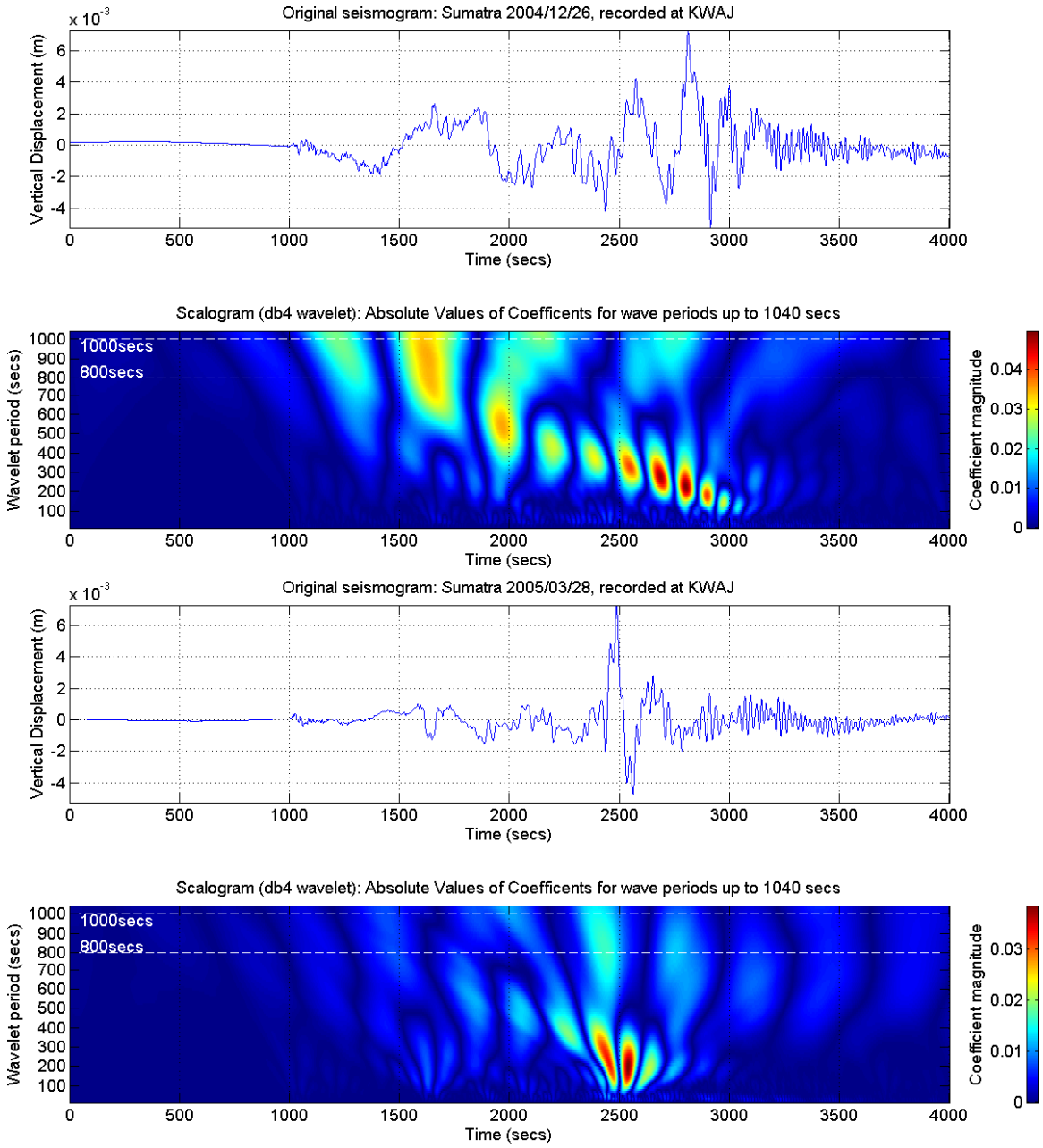

Figure 6. Seismograms and scalograms for KWAJ.

their approximately uniform distribution in azimuth. Clearly if closer stations could be used, then the early warning potential is greater; however, the signal-to-noise ratio of the records of the 2004 event at closer stations deteriorates and the data is as such not amenable to our analysis.

7. Conclusions and Further Work

[18] Wavelet analysis presents a useful way of observing the varying spectral (time-frequency) components of a seismogram. The wavelet method is able to simultaneously pick up both high- and low-frequency arrivals, as demonstrated in Figure 2b. Wavelet analysis is able to clearly distinguish the amplitudes of the long-period component between the seismograms of the 2004 Sumatra-Andaman earthquake and the 2005 Nias earthquake before the seismic S wave reaches the station. It shows that the 2004 earthquake generates a W phase of significantly greater amplitude. Due to the correspondence between strong long-period

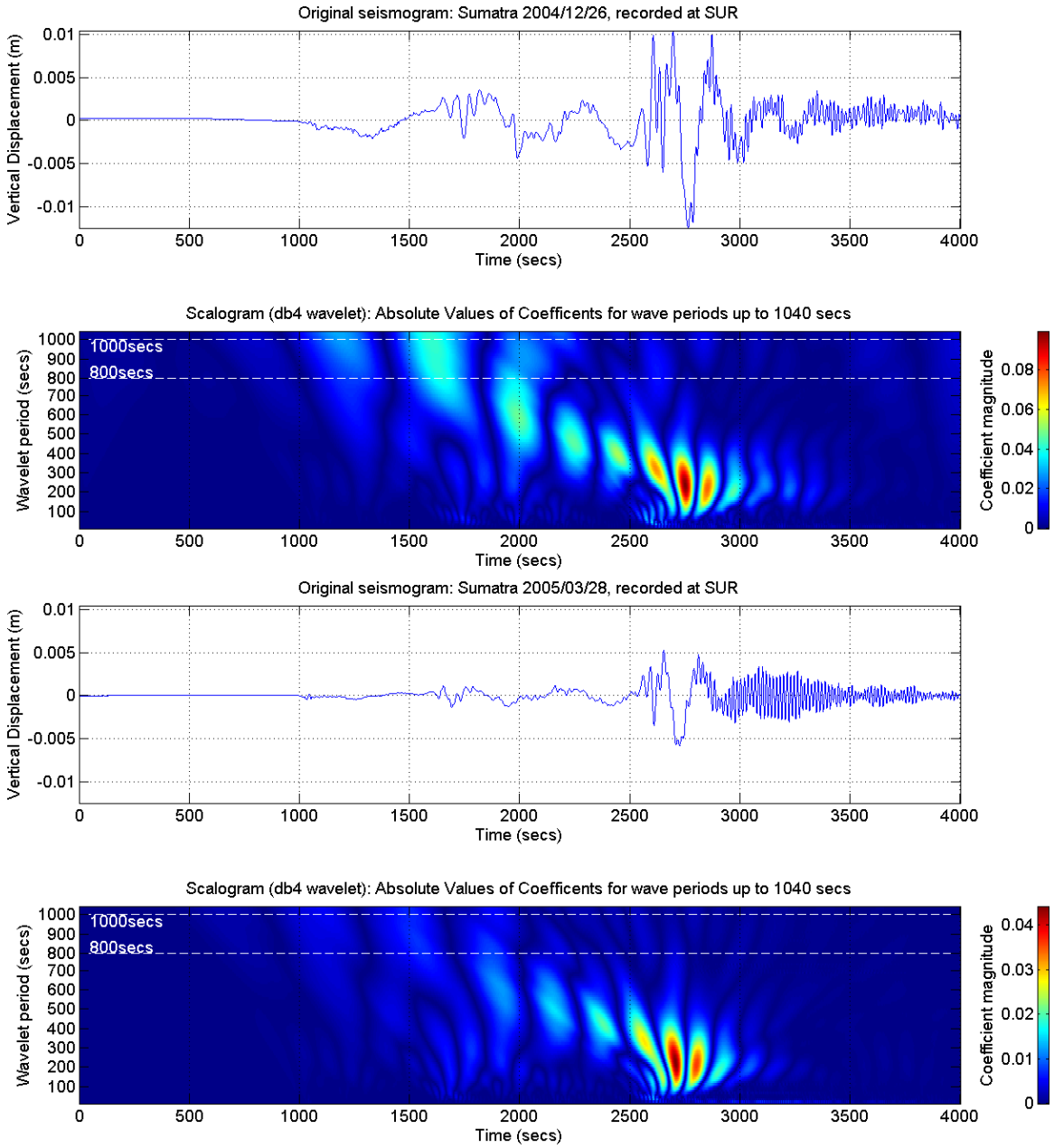

Figure 7. Seismograms and scalograms for SUR.

seismic components and tsunami generation, this facility has potential application for tsunami early warning.

[19] The W phase is produced by superposition of long-period (longer than 500 s) Rayleigh-wave overtones. Since these overtones have fairly constant group velocities, the superposition results in body-wave like appearance. In principle, the W phase should have distinct amplitude and phase spectra which can be utilized to identify it. However, we have not established a method to do this. To minimize false alarms, it would be desirable to distinguish W phase from spurious long-period waves caused by glitches or other noises.

[20] It would be worth looking at the specific example of "slow" or "tsunami earthquakes" [*Kanamori and Kikuchi*, 1993], by which we refer to relatively small earthquakes that have produced tsunami much greater than would normally be expected from their magnitude alone. This occurs especially, though not exclusively, when the surface-wave magnitude, M_s, is used. These

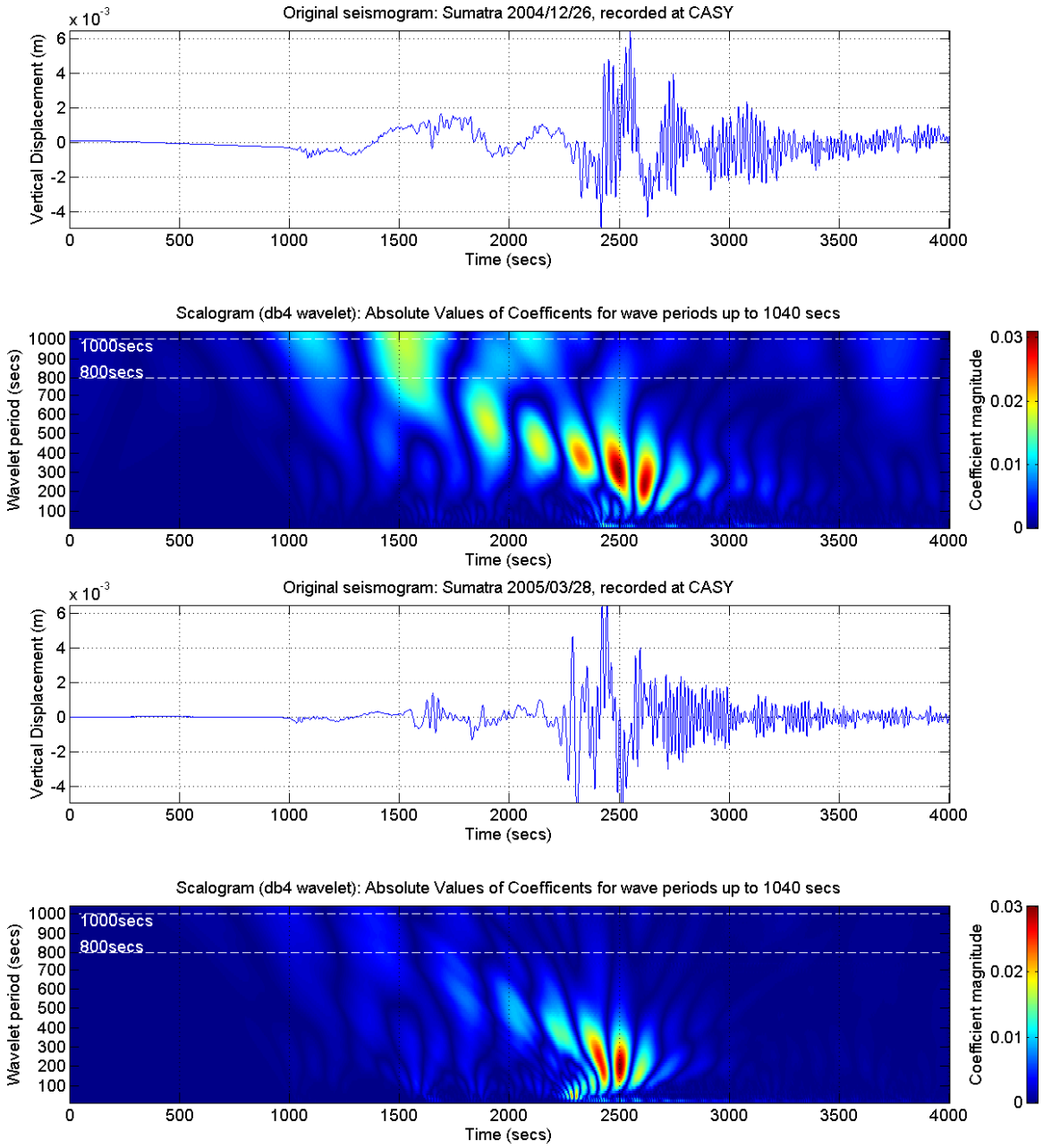

Figure 8. Seismograms and scalograms for CASY.

earthquakes, due to their slow rupture, tend to have a disproportionate amount of long-period energy. On this basis, we believe that our technique could recognize such events better than traditional tsunami detection methods.

[21] Overall the aim of an early warning implementation would be to set a threshold level of long-period coefficient amplitude (in terms of the wavelet-transformed recursive displacement seismogram) above which an oceanic subduction earthquake should generate a warning.

Acknowledgments

[22] The lead author is grateful to Caltech's SURF (Summer Undergraduate Research Fellowship) program for providing the opportunity to participate in this research and to L. Stolper for her invaluable support as organizer of the Cambridge-Caltech exchange. We used the Global Seismic Network seismograms of the Incorporated Research Institutions for

Seismology. We thank Meredith Nettles and an anonymous reviewer for their thoughtful comments.

References

Aki, K. (1972), Scaling law of earthquake source time function, *Geophys. J. R. Astron. Soc.*, *31*, 3–25.

Bormann, P., and K. Wylegalia (2005), Quick estimator of the size of great earthquakes, *Eos Trans. AGU*, *86*(46), 464.

Bryant, E. (2001), *Tsunami*, 320 pp., Cambridge Univ. Press, New York.

Chakraborty, A., and D. Okaya (1995), Frequency-time decomposition of seismic data using wavelet-based methods, *Geophysics*, *60*, 1906–1916.

Cohen, L. (1989), Time-frequency distributions: A review, *Proc. IEEE*, *77*(7), 941–979.

Huerta-Lopez, C., et al. (2000), Time-frequency analysis of earthquake records, paper presented at 12th World Conference on Earthquake Engineering, Int. Assoc. for Earthquake Eng., Auckland, New Zealand.

Kanamori, H. (1993), W phase, *Geophys. Res. Lett.*, *20*, 1691–1694.

Kanamori, H., and J. W. Given (1983), Use of long-period seismic waves for rapid evaluation of tsunami potential of large earthquakes, in *Tsunamis—Their Science and Engineering*, edited by K. I. a. T. Iwasaki, pp. 37–49, Terra Sci., Tokyo.

Kanamori, H., and M. Kikuchi (1993), The 1992 Nicaragua earthquake: A slow tsunami earthquake associated with subducted sediments, *Nature*, *361*, 714–716.

Kerr, R. A. (2005), South Asia tsunami: Failure to gauge the quake crippled the warning effort, *Science*, *307*, 201.

Lomax, A., and A. Michelini (2005), Rapid determination of earthquake size for hazard warning, *Eos Trans. AGU*, *86*(21), 202.

Menke, W., and V. Levin (2005), A strategy to rapidly determine the magnitude of great earthquakes, *Eos Trans. AGU*, *86*(19), 185, 189.

Misiti, M., Y. Misiti, G. Oppenheim, and J.-M. Poggi (1996), Wavelet toolbox user's guide (for MATLAB), MathWorks Inc., Natick, Mass.

National Oceanic and Atmospheric Administration (2006), Deep-ocean Assessment and Reporting of Tsunamis (DART), Natl. Buoy Data Cent., Stennis Space Center, Miss. (Available at http://www.pmel.noaa.gov/tsunami/Dart)

Okal, E. A., et al. (1991), Automatic estimation of tsunami risk following a distant earthquake using the mantle magnitude Mm, in *Proceedings of the 2nd UJNR Tsunami Workshop, Honolulu, Hawaii 5–6 November 1990*, pp. 229–238, Natl. Geophys. Data Cent., Boulder, Colo.

Talandier, J., and E. A. Okal (1989), An algorithm for automated tsunami warning in French Polynesia based on mantle magnitudes, *Bull. Seismol. Soc. Am.*, *79*, 1177–1193.

Zhu, L. (2003), Recovering permanent displacements from seismic records of the June 9, 1994 Bolivia deep earthquake, *Geophys. Res. Lett.*, *30*(14), 1740, doi:10.1029/2003GL017302.